SYMMETRY IN ELECTRODYNAMICS

From Special to General Relativity
Macro to Quantum Domains

Mendel Sachs
Department of Physics
State University of New York at Buffalo

This chapter is based on the part of my research program in general relativity on electromagnetism. Most of it has been published, since the late 1950s, and is referred to in the bibliography. The analysis of the theory of electromagnetism based on its underlying symmetry in relativity theory is re-assembled here with added discussion.

1. Introduction

An important lesson from Einstein's theory of relativity is that the underlying symmetry of any scientific theory reveals many far-reaching physical implications that are not obvious at first glance. In regard to the subject of electrodynamics and its unification with optics, the initially discovered relations in the 19th century, between electrical charges and their motions and the resulting electric and magnetic fields of force, led to a set of partial differential equations for the laws of electrodynamics. The formalism was not completed until Maxwell saw the need, ***based on symmetry***, for an extra term in the equation that relates current density to a resulting magnetic field. His addition of the extra term, called "displacement current", then yielded the full expression of "Maxwell's equations". The latter were recognized as the laws of electrodynamics, which were then seen to incorporate the laws of optics.

Indeed, it was Maxwell's generalization of the laws of electrodynamics that revealed that the radiation solutions of these equations, which would not have appeared in the earlier version (without the displacement current term) predicted all of the known optical phenomena. After Maxwell's investigation of these optical implications of electrodynamics, other portions of the spectrum of radiation solutions were predicted and discovered empirically: radio waves, X-rays, infrared radiation, gamma rays. Thus, it was Maxwell's intuitive feeling for the need of symmetry in his laws of electrodynamics that led to the full unification of electrodynamics and optics in the expression of Maxwell's equations.

James Clerk Maxwell died in 1879, the same year that Albert Einstein was born. Sixteen years later Einstein recognized that Maxwell's equations are *covariant* with respect to the Lorentz transformations between relatively moving inertial frames of reference i.e. reference frames that are in constant relative motion in a straight line. That is to say, Einstein recognized in 1895 that the laws of electrodynamics, expressed with Maxwell's field equations, must be in one-to-one correspondence in all possible inertial frames of reference, *from the view of any one of them.*[1]

The set of transformations of the spacetime coordinates that project the laws of electrodynamics from any observer's reference frame to any other (continuously connected) inertial frame such that the laws remain unchanged is the *symmetry group* of the theory of special relativity. It was discovered that this is the Poincare group.[2] Einstein then asserted, ten years later in 1905, that not only the laws of electrodynamics and optics, but *all* of the laws of nature must be covariant under the transformations of the symmetry group of relativity theory. This is the assertion of "the theory of special relativity." It is a statement of the objectivity of the laws of nature regarding all possible inertial frames of reference.[3]

In the next stage, Einstein generalized this symmetry rule to assert that all of the laws of nature must remain objective (covariant) with respect to transformations between frames of reference that are in arbitrary types of relative motion. This is the "theory of general relativity". As a first step, the theory of general relativity led to a new explanation for the phenomenon of gravitation, agreeing with the successful predictions of Newton's theory of universal gravitation, but predicting more effects not predicted by the classical theory. General Relativity thereby superceded Newton's law of universal gravitation.[4]

A significant feature of general relativity is the role of geometry in the mathematical representation of *all* of the laws of nature. For Einstein found that the Euclidean (flat) spacetime was not an adequate logic to represent the laws of interacting matter and radiation. Instead, he had to generalize to Riemannian (curved) spacetime geometry. The implication was that all of the laws of nature, including the laws of electrodynamics and optics, must be field laws that are mapped in such a curved spacetime. Later on, in his quest for a unified field theory, Einstein did insist that one should not only exploit the geometrical logic of the spacetime language, but one should also exploit its algebra. Here he referred to the underlying group of general relativity – which I have called "the Einstein group". In a 1945 article,[5] Einstein said: "Every attempt to establish a unified field theory must start, in my opinion, from the group of transformations which is no less general than that of the continuous transformations of the four coordinates".

At the outset, then, it is important to recognize in our study of a generalization of the laws of electrodynamics based on the full symmetry of relativity theory that its covariance is in terms of a *continuous group*[6] (whether we refer to special relativity or to general relativity). Such a group does not admit the discrete reflections in space or time. Further, because of the requirement of incorporating laws of conservation of energy, momentum and angular momentum, in the flat spacetime limit of the theory, *Noether's theorem* prescribes that the transformations that define the covariance in relativity theory must be analytic.[7] That is to say, the relativistically covariant solutions of the laws of nature must be regular (i.e. nonsingular) *everywhere*.[8] Such groups of continuous, analytic transformations (the Poincare group for special relativity[2] and the Einstein group for general relativity) are *Lie groups*.[9] They prescribe the algebraic logic of the theory of relativity.

In Section 2 there will be an outline of the generalization of the vector potential of electromagnetic theory so as to include a gauge invariant *pseudovector* part. This is allowed because of the lack of reflection symmetry in the relativity groups. In Section 3 the full form of the equations of electrodynamics in terms of the irreducible representations of the Lie groups of relativity theory will be shown. It is a two-component spinor formalism that follows from a factorization of the standard vector representation of the Maxwell formalism. In Section 4 the theory will be extended to its full form in general relativity. It will be shown that the 16-component quaternion metrical field equation emerges as a factorization of Einstein's (10-component) symmetric tensor field equations in general relativity, once the reflection symmetry elements are removed from the underlying covariance group. It is then demonstrated that these 16 independent field equations may be re-written as a sum of 10 second -rank symmetric tensor equations, corresponding exactly with the Einstein field equations, plus 6 second-rank antisymmteric tensor equations. It is shown that the latter may be put into a form that corresponds exactly with the formal structure of Maxwell's equations. Thus, it is because of removing the reflection symmetry elements from the underlying group of general relativity, that one arrives at the factorized field equations that fully unify the gravitational features of matter, in terms of Einstein's field equations, with the electromagnetic features of matter in terms of the Maxwell field equations. The *route* toward achieving a unified field theory from general relativity is then to follow the rules of the underlying Lie group by removing the reflection symmetry elements from the symmetry group of Einstein's tensor field equations, thereby yielding a natural structure of the formalism in terms of spinor and quaternion variables. It will be seen in this analysis that, in accordance with a *generalized Mach principle*,[10] the electromagnetic field of a charged body vanishes in a vacuum.

In Section 5 it will be shown that the quaternion structure of the fields that correspond with the electromagnetic field tensor and its current density source, implies a very important consequence for electromagnetism. It is that the local limit of the time component of the 4-current density yields a *derived* normalization. The latter is the condition that was *imposed* (originally by Max Born) to interpret Quantum Mechanics as a probability calculus. Here, it is a derived result that is an asymptotic feature (in the flat spacetime limit) of a field theory *that may not generally be interpreted in terms of probabilities.* Thus, the derivation of the electromagnetic field equations in general relativity reveals, as a bonus, a natural normalization condition that is conventionally imposed in quantum mechanics.

2. A Generalization of the Electromagnetic Potential

After the momentous discovery by C.S. Wu and her collaborators in 1957 that the weak interaction violates parity (spatial reflection),[11] I addressed the question on whether there may be any empirical evidence for the violation of parity in the electromagnetic interaction.[12] It was thought that only the weak interaction violates space-reflection symmetry. But if, in the final analysis, there is a unified field theory in which the weak and the electromagnetic forces, as well as all of the other forces of nature, are manifestations of a single force field, then there is an implication in the experimental result of Wu *et al* that the electromagnetic and the nuclear forces also violate space-reflection symmetry, as well as time reversal symmetry. Indeed, the ***continuous group*** that underlies relativity theory implies that *all* discrete symmetries must be excluded from the laws of nature.

At that time, in the 1950s, there was a problem whereby the calculations from quantum electrodynamics for the Lamb shift, $2S_{1/2} - 2P_{1/2}$ in the states of hydrogen, were not in exact agreement with the measurements. Thus it occurred to me that a small violation of parity symmetry in the electromagnetic interaction might be responsible for this discrepancy.

My investigation proceeded along the following line:[12] The conventional coupling of an external electromagnetic potential field A_μ to an electrical four-current density of matter j^μ is in terms of the *scalar* interaction Lagrangian,

$$L_{int}^{(s)} = j^\mu A_\mu = e\psi^{(e)+}\gamma^0\gamma^\mu\psi^{(e)}A_\mu \qquad (2.1)$$

$\psi^{(e)}$ is the (four-component) bispinor electron field that solves the Dirac equation

$$[\gamma^\mu(\partial_\mu + ieA_\mu) + m]\psi^{(e)} = 0 \qquad (2.2)$$

It is assumed here that the electromagnetic vector potential A_μ is a (polar) four-vector field. Thus the Lagrangian $L_{int}^{(s)}$ is a scalar function in space and time.

If parity should be violated, A_μ may be generalized by adding an (axial) pseudovector part, B_μ to A_μ. The Lagrangian then generalizes to a sum of a scalar part and a pseudoscalar part, $L_{int} = L_{int}^{(s)} + L_{int}^{(ps)}$, where the latter part , $e\psi^{+}\gamma^{0}\gamma^{\mu}\psi B_\mu$, is clearly a pseudoscalar function.

Then what is the source of B_μ in electromagnetic theory? Are there restrictions on A_μ that should also apply to B_μ? The answer is yes – it is the restriction of gauge invariance in order to yield a unique representation for the electric and magnetic field variables. Additionally, gauge invariance is the necessary and sufficient condition for the existence of conservation laws in the formalism – in this case the requirement of the conservation of electrical charge.[13] The latter follows from the continuity equation,

$$\partial^{\mu} j_\mu = 0. \qquad (2.3)$$

The gauge covariance is in two parts: l) gauge invariance of the first kind, that is, invariance of the formalism under the phase change $\psi \rightarrow \psi\exp(i\eta)$, where, at this stage, $\eta(x)$ is an arbitrary scalar field, and 2) gauge invariance of the second kind - applied to the vector potential, this is the change: $A_\mu \rightarrow A_\mu + ie\partial_\mu\eta$. The latter is the addition of a four-gradient of the scalar field η to the original potential A_μ. Applying these two types of gauge transformations to the Dirac Equation (2.1) then leaves it covariant, i.e. its form is left unchanged.

The Unique Form of B_μ

The form of the pseudovector potential B_μ that is gauge invariant would be a field that effectively interchanges the roles of the electric and magnetic field variables in the interaction Hamiltonian that couples a charge q to external electric and magnetic fields. For such a form leaves the electromagnetic interaction with charged matter still dependent only on the electric and magnetic field variables, directly.

The idea is the following: The electromagnetic field intensity solution of Maxwell's equations, expressed in terms of the four-dimensional curl of the vector potential is:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \qquad (2.4)$$

where the antisymmetric second-rank tensor $F_{\mu\nu}$ is the combination of the electric and magnetic field variables, **E** and **H**, as follows: $F_{0k} = -F_{k0} = E_k$, $F_{jk} = -F_{kj} = H_n$ and $F_{\mu\mu} = 0$., where $j \neq k \neq n = 1, 2, 3$ and the '0' subscript is the time component. $F_{\mu\nu}$ are the solutions of Maxwell's equations:

$$\partial^{\nu} F_{\mu\nu} = 4\pi j_\mu\,, \qquad \partial_{[\rho,}F_{\mu\nu]} = 0 \qquad (2.5)$$

The 'bracket' in the second equation in (2.5) denotes a cyclic sum and we use units (henceforth in this article) with $c = 1$. Combining the definition of $F_{\mu\nu}$ as the four-dimensional curl of a four vector, as in eq. (2.4), Maxwell's equations in terms of the vector potential are:

$$\Box A_\mu = 4\pi j_\mu \qquad (2.6)$$

where $\Box$ is the D'Alembertian operator $\partial^2/\partial t^2 - \nabla^2$, and the vector potential is subject to the *Lorentz gauge,* $\partial^\mu A_\mu = 0$,
which in turn corresponds to selecting the phase $\eta(x)$ to be a solution of the wave equation, $\Box\eta = 0$.

Since j_μ is a four-vector field and $\Box$ is a scalar operator, it follows that A_μ is a (polar)vector field. Let us now choose the (axial) pseudovector field B_μ that accompanies A_μ so that 1) it satisfies the same Lorentz gauge as A_μ, i.e. $\partial^\mu B_\mu = 0$, and it solves the field equation (that accompanies (2.6) for A_μ):

$$\Box B_\mu = -i\xi(4\pi/2)\varepsilon_{\mu\nu\lambda\rho}\int\partial_\rho j_\lambda dx_\mu \qquad (2.7)$$

where $\varepsilon_{\mu\nu\lambda\rho}$ is the totally antisymmetric Levi-Civita symbol, with $\varepsilon_{0123} = +1$. ξ is an undertermined parameter at this stage of the analysis, it is, physically, a measure of the ratio of reflection-nonsymmetric terms to the reflection-symmetric terms in the generalized four-potential expression of electrodynamics, as discussed above.

A solution of eq. (2.7) that is compatible with the continuity equation, $\partial^\mu j_\mu = 0$, which in turn leads to the conservation of charge, is:

$$B_\nu = -(i\xi/2)\int\varepsilon_{\mu\nu\lambda\rho}\partial_\rho A_\lambda\, dx_\mu \qquad (2.8)$$

The integration above is defined to be an *indefinite line integral*. [It is, by definition, a function of the spacetime coordinates x, not dependent on limits in the line integral. For example, the line integral $\int x\, dx$ is *defined* here to be $x^2/2$.] The integrand in eq. (2.8) is the dual of the electromagnetic field tensor, $F_{\mu\nu} = \varepsilon_{\mu\nu\lambda\rho}F_{\lambda\rho}$. That is, it replaces the electric and magnetic field variables, $\mathbf{E} \longleftrightarrow \mathbf{H}$. In terms of these variables, the pseudovector potential may be expressed as follows: $B_\nu = \{\mathbf{B}, B_0\}$, where

$$\mathbf{B} = (\xi/2)\int(\mathbf{H}dt + \mathbf{E} \times d\mathbf{r}), \quad \mathbf{B}_0 = (i\xi/2)\int\mathbf{H}\bullet d\mathbf{r} \qquad (2.9)$$

The Case of Constant Fields

If **E** and **H** are constant fields (i.e. independent of the space and time coordinates), then they would come out of the integral signs in eq. (2.9). With the reflection non-symmetric Lagrangian density $L_{int} = j^\mu(A_\mu + B_\mu)$, Lagrange's equation of motion then reduces to the following equation of motion of a test body with charge q:

$$d\mathbf{p}/dt = \mathbf{F}_v + \mathbf{F}_{pv}$$

where **p** is the particle's momentum,

$$\mathbf{F}_v = q[\mathbf{E} + \mathbf{v} \times \mathbf{H}] \qquad (2.10)$$

is the usual vector (polar) Lorentz force in electrodynamics and

$$\mathbf{F}_{pv} = -q\,\xi[\mathbf{H} - \mathbf{v} \times \mathbf{E}] \qquad (2.11)$$

is a pseudovector (axial) contribution, which I have called the "anti-Lorentz force". The latter predicts that a charge q would move along the lines of the magnetic field **H.** Even if the value of the parameter ξ should be extremely small, a very large magnitude of the magnetic field intensity **H,** say in the interior or near a rotating galaxy, may make this prediction observable in astrophysical measurements. The second term in F_{pv} predicts a motion of a charge q in an external electric field **E** such that it would rotate perpendicularly to the plane of its velocity vector **v** and the imposed constant electric field **E**.

The Generalized Dirac Hamiltonian

The behavior of an electron in an electromagnetic field, in the context of the quantum theory, is determined from the solutions of the Dirac equation. Here the free particle momentum operator is replaced with the generalized four-momentum operator, $p_v + e(A_v + B_v)$. The Dirac equation then takes the form:

$$\{\gamma^v(p_v + e(A_v + B_v) - im\}\psi = 0 \qquad (2.12)$$

where $p_v = -i\partial_v$ (units are used with $h/2\pi = 1$) and the 'Dirac matrices' are:

$$\gamma^v = \{\gamma, \beta\}, \quad \gamma = -i\beta\alpha, \quad \gamma^4 = \beta$$

and α, β are the 4 x 4 matrices defined in terms of the Pauli matrices: $\alpha_{12} = \alpha_{21} = \sigma, \alpha_{11} = \alpha_{22} = 0, \quad \beta_{11} = -\beta_{22} = I, \beta_{12} = \beta_{21} = 0$, I is the unit 2-matrix and

$$\gamma_\mu \gamma_v + \gamma_v\gamma_\mu = 2\delta_{\mu v} \qquad (\mu, v = 1, 2, 3, 4)$$

The generalized Dirac equation was applied to the case of the hydrogen atom.[14] It was to investigate whether the added potential B_μ in the Dirac Hamiltonian in eq. (2.12) would predict a contribution to the Lamb shift, exclusive of quantum electrodynamics. The exact solutions of (2.12) were determined for the hydrogen atom. The very interesting (and unexpected!) result was found that the added potential did not lift the accidental degeneracy in the states of hydrogen. That is, there was no prediction of any contribution to the Lamb shift from B_μ that might have accounted for the small difference between the experimental observations and the predictions of quantum electrodynamics (in the late 1950s).

The pseudovector four-potential B_μ may still contribute to other effects in the microscopic domain. For example, it would predict that a particle, such as a neutron, would have an electric dipole moment, whose value is proportional to the term in the Dirac Hamiltonian $\xi\sigma\bullet E$.[12] However, after much experimental investigation into the possibility of the neutron electric dipole moment, it has not been found[15] – that is, in the context of this theory, the parameter ξ, if it were non-zero, must be too small (the order of 10^{-13}) for this effect to be observed.

A later analysis was based on a spinor formulation of the electromagnetic field theory, to be discussed in the next section. It was found that this generalization, based on fully conforming to the reflection non-symmetric field theory that is in accordance with the symmetry group of relativity theory, the Lamb shift is indeed fully predicted, exclusive of quantum electrodynamics. The predictions were in agreement with the empirical facts, within the experimental error, for the hydrogen states with principal quantum numbers n = 2, 3, 4.[16]

3. Factorization of Maxwell's Equations to a Spinor Form

In the context of Einstein's theory of relativity, it must be asked: Is Maxwell's expression of the electromagnetic theory the most general representation consistent with the symmetry requirements of relativity? The answer is negative because the symmetry of Maxwell's equations based on *reducible* representations of the group of relativity theory. Then there must be additional physical predictions that remain hidden that would not be revealed until the most general (*irreducible*) expression of the electromagnetic field theory is used.

Two equivalent forms of Maxwell's field equations in terms of the standard vector formalism are eq. (2.5), or (2.6) with the Lorentz gauge $\partial^{\mu}A_{\mu} = 0$. The former is in terms of the antisymmetric second rank tensor solution $F_{\mu\nu}$, that is a combination of the electric and magnetic field variables. The latter is in terms of the vector potential, A_{μ}, shown in eq. 2.6 (as well as eq. (2.7) in terms of the pseudovector potential B_{μ}, assuming that the parameter ξ is nonzero). [Experimental results to this point in time indicate that indeed this parameter is zero to within experimental accuracy[15] – even though the symmetry of relativity theory has no reason to exclude it. Henceforth, we will assume that this parameter is zero.]

The symmetry requirements of the theory of relativity have geometrical and algebraic modes of expression. From the geometrical view in special relativity, the continuous spacetime transformations that leave the laws of nature covariant (i.e. unchanged in form) in all possible inertial frames of reference, *from the view of any one of them*, are the same set of transformations that leave invariant the squared differential metric

$$ds^2 = (dx^0)^2 - dr^2 \qquad (3.1)$$

In general relativity, where the relative motion between frames is not inertial, the geometrical invariant of the resulting 'curved' spacetime is

$$ds^2 = g^{\mu\nu}(x)dx_\mu \, dx_\nu \qquad (3.2)$$

where μ and ν are summed from 0 to 3 and $g^{\mu\nu} = g^{\nu\mu}$ is the ten-component metric tensor with the flat spacetime limit that takes eq. 3.2 into 3.1.That is, in the local limit of a flat spacetime,

$$g^{\mu\nu}(x) \rightarrow (g^{00} = 1, g^{kk} = -1 \ (k = 1, 2, 3) \text{ and } g^{\mu\neq\nu} = 0)$$

The idea of covariance is then that the same set of spacetime transformations that leave the differential metric (3.1) in special relativity, or (3.2) in general relativity, unchanged (invariant) also leave all of the laws of nature covariant (unchanged in form) under these transformations between reference frames. The metric (3.1) in special relativity, or (3.2) in general relativity, then guides one to the forms of the covariant laws of nature, in accordance with the theory of (special or general) relativity. This is the role of the differential metrics – *they are not to be considered as 'observables' on their own*!

A significant point here is that it is not the *squared invariant* ds^2 that is to underlie the covariance of the laws of nature. It is rather the *linear invariant* ds that plays this role. Then how do we proceed from the squared metric to the linear metric? That is to say, how does one take the 'square root' of ds^2? The answer can be seen in Dirac's procedure, when he factorized the Klein-Gordon equation to yield the spinor form of the electron equation in wave mechanics – i.e. the 'Dirac equation'. Indeed, Dirac's result indicated that by properly taking the 'square root' of ds^2 in relativity theory, extra spin degrees of freedom are revealed that were previously *masked.*

The symmetry group of relativity theory tells the story. For the *irreducible representations* of the Poincare group (of special relativity) or the Einstein group (of general relativity) obey the algebra of quaternions. The basis functions of the quaternions, in turn, are two-component spinor variables.[17]

We start out then with a factorized metric in special relativity, which has the quaternion form:

$$ds = \sigma^\mu dx_\mu \equiv \sigma^0 dx_0 - \boldsymbol{\sigma} \bullet d\mathbf{r} \qquad (3.3)$$

where σ^0 is the unit two-dimensional matrix and σ^k ($k = 1, 2, 3$) are the three Pauli matrices. The set $\{\sigma^\mu\}$ form the four basis elements of a quaternion (analogous to the two basis elements $\{1, i\}$ of a complex number).

In the global extension to general relativity, the geometric generalization from the flat spacetime description to a curved spacetime, the basis elements $\sigma^\mu \rightarrow q^\mu(x)$, so that the factorized invariant differential element becomes

$$ds = q^\mu(x)dx_\mu \qquad (3.4)$$

This quaternion differential is a generalization of the Riemannian metric. The four-vector *quaternion fields* $q^\mu(x)$ then replace the second-rank, symmetric tensor fields $g^{\mu\nu}(x)$ as the fundamental metric of the spacetime. The metric field $q^\mu(x)$ is a four-vector, whose four components are each quaternion- valued. This is then a 16-component field, rather than the 10-component metric tensor field $g^{\mu\nu}$ of the standard Riemannian form.

The 16-component quaternion metric field is then a generalization of the 10-component metric tensor field to represent gravitation. The increase in the number of components satisfies the group requirement of general relativity theory – that the Einstein group is a 16-parameter Lie group, indicating that there must be 16 essential parameters to characterize the irreducible representations of the group. This implies that there must be 16 independent field equations to underlie the spacetime metric. These are the 16 essential parameters of the Einstein group. They are the derivatives of four coordinates $x^{\mu'}(x)$ of one reference frame with respect to those of another: $\partial x^{\mu'}/\partial x^\nu$ $(\mu,\nu = 0, 1, 2, 3)$ The basic reason for the increase in the number of components of the metric field q^μ, compared with $g^{\mu\nu}$, is that the reflection symmetry elements of the spacetime have been removed from the underlying symmetry group of the latter. [It is the same reason that the removal of the reflection symmetry elements from the covariance of the Klein-Gordon operator yields the extra (spinor) degrees of freedom in the factorized Dirac operator in quantum mechanics]. Thus, the factorized metric (3.4) has no reflection symmetry while the 'squared' metric in special relativity (3.2) does have reflection symmetry in space and time.

The 'key' to the generalization achieved is then the removal of the reflection symmetry elements in the space and time coordinates in the laws of nature. This then leads to the Poincare group (of special relativity) or the Einstein group (of general relativity), since these are Lie groups – groups of only continuous, analytic transformations of the spacetime coordinate systems that leave the laws of nature covariant.

Let us now focus on the irreducible expressions of the electromagnetic field equations in special relativity, using the quaternion calculus. We will then come to their form in general relativity.

Following from the quaternion differential metric (3.4), we have the first order quaternion differential operator:

$$\sigma^\mu \partial_\mu = \sigma^0 \partial_0 - \sigma \bullet \nabla$$

The basis functions of this operator are the two-component spinor variables. Guided by the two-dimensional Hermitian structure of the representations of the Poincare group, we may make the following identification between the spinor basis functions ϕ_α ($\alpha = 1,2$) of this operator and the components (E_k, H_k) ($k = 1, 2, 3$) of the electric and magnetic fields, in any particular Lorentz frame:

$$(\phi_1)_1 = G_3\,,\ (\phi_1)_2 = G_1 + iG_2\,,\ (\phi_2)_1 = G_1 - iG_2\,,\ (\phi_2)_2 = -G_3$$
$$(\Upsilon_1)_1 = -4\pi i(\rho + j_3)\quad (\Upsilon_1)_2 = -4\pi i(j_1 + ij_2)\quad (\Upsilon_2)_1 = -4\pi i(j_1 - ij_2)$$
$$(\Upsilon_2)_2 = -4\pi i(\rho - j_3) \qquad (3.5)$$

where $G_k = H_k + iE_k$. It is then readily verified that the two uncoupled, two-component spinor equations:

$$\sigma^\mu \partial_\mu \phi_\alpha = \Upsilon_\alpha \quad (\alpha = 1, 2) \qquad (3.5')$$

precisely duplicate the standard form (2.5) of Maxwell's equations.[18]

It is important to note at this stage of the analysis that the generalization achieved in going from the vector representation (2.5) to the spinor representation (3.5') of the electromagnetic field equations is not merely a re-writing of the Maxwell equations. This is because the spinor formalism has more degrees of freedom than the vector formalism, thus it makes more predictions, in addition to duplicating the predictions of the (less general) vector form of the theory. This will be demonstrated in the following paragraphs.

Under the Poincare group of transformations of special relativity, when $x^\mu \rightarrow x^{\mu\prime} = \alpha^\mu{}_\nu x^\nu$, where $\{\alpha^\mu{}_\nu\}$ are the vector transformations, covariance of the spinor field equations (3.5) is preserved *if and only if*[17]

$$\phi_\alpha(x) \rightarrow \phi_\alpha'(x') = S\phi_\alpha(x)\,,\quad \Upsilon_\alpha(x) \rightarrow \Upsilon_\alpha'(x') = S^{+-1}\Upsilon_\alpha(x) \qquad (3.6)$$

where the spinor transformations S relate to the vector transformations $\alpha^\mu{}_\nu$ according to the equation:

$$S^+\sigma^\mu S = \alpha^\mu{}_\nu \sigma^\nu \qquad (3.7)$$

Equation 3.7 then yields the double-valued spinor transformation:

$$S(\theta_{\mu\nu}) = \exp(\sigma^\mu\sigma^\nu\theta_{\mu\nu}/2) \qquad (\mu,\nu = 0, 1, 2, 3) \qquad (3.8)$$

Note that this equation is not summed over (μ,ν). $\theta_{\mu\nu}$ are the constant (i.e. x-independent) parameters the define the ten transformations in the x_μ-x_ν plane and the four translations of the 10-parameter Poincare group: three Eulerian angles of rotation in space, three components of the relative speed between inertial frames, and the four translations in space and time.

The solutions $F_{\mu\nu}$ of the standard (reducible) vector form of electromagnetic field theory transform as a second-rank (covariant) tensor

$$x \rightarrow x' \Rightarrow F_{\mu\nu}(x) \rightarrow F_{\mu\nu}'(x') = \alpha^\lambda{}_\mu \alpha^\rho{}_\nu F_{\lambda\rho}(x) \qquad (3.9)$$

Thus the *identification* (3.5) $\phi_\alpha(F_{\mu\nu})$ is not to be understood as form-invariant regarding the dependence (3.5) of the spinor variables ϕ_α on the tensor variables $F_{\mu\nu}$ in any other Lorentz frame. That is to say, the Lorentz transformation of $\phi_\alpha(F_{\mu\nu})$ does not transform form-invariantly into $\phi_\alpha'(F_{\mu\nu}')$ under the Lorentz transformations of the Poincare group, $x \rightarrow x'$. Of course, this is because ϕ_α transforms as a spin one-half basis function of the irreducible representations of the group of relativity while $F_{\mu\nu}$ transforms as the basis functions of the spin-one (reducible) representations of the group.

The terms of the respective (reducible) tensor and the (irreducible) spinor expressions of the electromagnetic laws that must correspond in all Lorentz frames are those that identify with physical observations. These are the conservation laws of electromagnetism. They derive, in turn, from the invariants of the theory.

In accordance with the transformation properties (3.6), it follows that the Hermitian products,

$$I_{\alpha\beta} = \phi_\alpha^+\Upsilon_\beta \qquad (\alpha,\beta = 1 \text{ or } 2) \qquad (3.10)$$

are four complex number invariants, thus corresponding to eight real number invariants. Particular linear combinations of these invariants may then be set up to correspond with the standard invariants of the vector form of electromagnetic theory. Since there are more independent invariants here than in the standard theory, there must be more invariants and physical predictions that have no counterpart in the standard form of the theory.

According to the spinor calculus,[19] further invariants, in addition to (3.10), that correspond with the standard invariants of electromagnetic field theory, are:

$$I_1 = \phi_1^{tr}\varepsilon\phi_2 \Leftrightarrow (E^2 - H^2) + 2i\mathbf{E}\bullet\mathbf{H} \qquad (3.11a)$$

$$I_2 = \Upsilon_1^{tr}\varepsilon\Upsilon_2 \Leftrightarrow \rho^2 - j^2 \qquad (3.11b)$$

where 'tr' stands for the transpose of the spinor variable and ε is the two-dimensional Levi-Civita symbol, with $\varepsilon_{01} = -\varepsilon_{10} = 1$, and $\varepsilon_{00} = \varepsilon_{11} = 0$.

We see here that the real and oimaginary parts of the complex invariant I_1 corresponds with the two invariants of the standard form of electromagnetic theory, the scalar and the pseudoscalar terms. They appear together here in a single complex function because of the reflection-nonsymmetric feature of this theory. The invariant I_2 corresponds with the real-valued modulus of the 4-current density of the standard theory.

The Conservation Equations

It follows from the spinor field equation (3.5') that these equations may be re-written in the form of four complex conservation equations:

$$\partial_\mu(\phi_\alpha^+\sigma^\mu\varphi_\beta) = \phi_\alpha^+\Upsilon_\beta + \Upsilon_\alpha^+\phi_\beta \qquad (3.12)$$

If we set $\alpha = \beta = 1$ in (3.12) and add this to the equation with $\alpha = \beta = 2$, it follows from the identification (3.5) that their sum corresponds with the standard form of the conservation equation:

$$(1/2)\partial^0(E^2 + H^2) + \mathrm{div}(\mathbf{E} \times \mathbf{H}) = -4\pi\, \mathbf{E}\bullet\mathbf{j} \qquad (3.13a)$$

$$\partial^0(\mathbf{E} \times \mathbf{H}) = \rho\mathbf{E} + \mathrm{j} \times \mathbf{H} \qquad (3.13b)$$

Thus we see that four of the conservation equations in (3.12) correspond with all of the four conservation equations of the standard theory: One is the conservation of energy (3.13a) (Poynting's equation), and the other three are the conservation of the three components of momentum (3.13b) of the standard form of electromagnetic field theory. But since (3.12) are eight real-number valued equations rather than four, the spinor formalism predicts more facts than the standard vector Maxwell formalism – it is a true generalization.

Faraday's Interpretation

With Faraday's interpretation of the electromagnetic field as a 'potentiality' of force exerted by charged matter, then to be actualized by a test body at the spacetime point x where it is located, there must be a separate field of force for each charged source. Thus, Maxwell's equations (2.5) must be labeled for each source field:

$$\partial^\nu F^{(n)}{}_{\mu\nu} = 4\pi j^{(n)}{}_\mu \qquad \partial_{[\rho} F^{(n)}{}_{\mu\nu]} = 0 \qquad (3.14)$$

for the nth source field. Similarly, the spinor expression of the electromagnetic equations are:

$$\sigma^\mu \partial_\mu \phi_\alpha{}^{(n)} = \Upsilon_\alpha{}^{(n)} \qquad (3.15)$$

It is important to note that while there are fields for each of the sources of the system, they are all mapped in the same spacetime x, rather than separate spacetimes for each source. This is the ***nonlocal*** feature of this field theory since there are no individual trajectories for charged discrete particles. This interpretation eliminates the problem of the self-energy of the electron, as a singular, charged particle of matter.[20]

The conserved energy in the vector representation with Faraday's interpretation is then:

$$\Sigma_n \Sigma_{m \neq n} (1/16\pi)(\mathbf{E}^{(m)}\bullet\mathbf{E}^{(n)} + \mathbf{H}^{(m)}\bullet\mathbf{H}^{(n)}) \qquad (3.16)$$

rather than the standard form

$$(1/8\pi)(E^2 + H^2) \qquad (3.17)$$

The form (3.17) is derived from integrating the conservation law (3.13a) over all of space and using Gauss' law. It includes the self-energy terms (m = n) as well as the free field (radiation) terms that are independent of any sources. The latter terms are automatically absent from the expression (3.16) – which is finite from the outset and entails no 'free radiation'.

In the spinor formalism (3.15), the four *complex* conservation equations, with Faraday's interpretation are:

$$\partial_\mu \Sigma_n \Sigma_{m\neq n} \phi_\alpha^{(n)+} \sigma^\mu \phi_\beta^{(m)} = \Sigma_n \Sigma_{m\neq n} (\phi_\alpha^{(n)+} \Upsilon_\beta^{(m)} + \Upsilon_\alpha^{(n)+} \phi_\beta^{(m)}) \qquad (3.18)$$

The right-hand side of this scalar equation, which are four complex relations, then entails eight real number scalar equations. As in the vector formalism, there are no self-energy terms present.

It is to be noted that some of the eight equations may be expressed in one-to-one correspondence with all of the four conservation equations of the standard Maxwell formalism. But there are other conservation equations here that have no counterpart in the standard formalism of electromagnetism. It further implies that indeed this is a true generalization of the Maxwell form of electromagnetism.

It is interesting to note here a difference between the standard theory and the Faraday interpretation that relies on the Mach principle. Consider the complex conservation equation (3.18) with $\alpha = \beta = 1$. The imaginary part of this complex equation is:

$$\partial_\mu \Sigma_n \Sigma_{m\neq n} (\phi_1^{(m)+} \sigma^\mu \phi_1^{(n)} - \phi_1^{(n)+} \sigma^\mu \phi_1^{(n)}) = \Sigma_n \Sigma_{m\neq n} (\phi_1^{(m)+} \Upsilon_1^{(n)} - \Upsilon_1^{(n)+} \varphi_1^{(m)})$$
$$+ (\Upsilon_1^{(m)+} \phi_1^{(n)} - \phi_1^{(n)+} \Upsilon_1^{(m)})$$

If m = n, as included in the standard Maxwell theory, the extra four conservation equations above reduce to 0 = 0 – an ambiguity. However, with the restriction from Faraday's interpretation that requires that $m \neq n$, the ambiguity is removed and the extra conservation equations remain.

Let us now sum up the generalization of electromagnetic field theory thus far. The starting point is that the symmetry group that underlies Einstein's theory of relativity is a Lie group – a group of continuous, analytic transformations that preserve the covariance of all of the laws of nature. This is the rule that all of the laws of nature remain in one-to-one correspondence in all continuously connected reference frames, *from the view of any one of them.* This group does not entail any discrete transformations in space or time.

In the first section of this paper it was shown that the only gauge invariant way to express this extension in the context of the vector potential expression of the electromagnetic field theory is to add to the standard (polar) vector potential A_μ an (axial) pseudovector contribution B_μ. This has the effect of interchanging the roles of the electric and the magnetic field variables in the coupling of a charged body to this field. While this addition is theoretically permitted, it was found to not imply any significant empirical contributions, for the magnitudes of fields studied thus far in the experimental domain. Future experimental studies of very high magnetic fields, such as the interiors of galaxies, may reveal possible observable consequences in astrophysical studies.

In the second section of this paper, the group requirement that removes the reflection transformations was met head-on, without the need to add any new potential terms. It was seen that the most general expression of electromagnetic field theory – that excludes reflections in spacetime in the underlying symmetry group, follows from a factorization of the vector representation of the Maxwell theory to a first-rank spinor form.
This is a natural generalization, following from the *irreducible representations* of the covariance group of relativity theory that leads to extra physical predictions, because of the extra degrees of freedom in the spinor variables.

In the next section, the final generalization of full symmetrization will be carried out, whereby the spinor-quaternion expression of the laws of electromagnetism will be extended from special relativity to general relativity. This extension automatically fuses the laws of electromagnetism with those of gravity. It will be shown that the generalized formalism for electromagnetism is obtained from a factorization of Einstein's field equations. The new formalism is in terms of a replacement of Einstein's tensor metric field with a vector field in which each of its four components is quaternion-valued. Thus, the new metric variable $q^\mu(x)$ has 16 independent components, rather than the 10 (of the symmetric tensor $g^{\mu\nu}$ of the Einstein formalism) or the 6 (of the antisymmetric tensor $F^{\mu\nu}$ of the Maxwell formalism). From the factorized metrical field equations in q^μ it will be seen how the Einstein formalism and the Maxwell formalism are recovered, though now identifying each with the single quaternion field and its related spinor calculus in a curved spacetime.

4. Extension of Electromagnetic Field Theory to General Relativity

The Group

In accordance with the principle of general covariance – the underlying axiom of the theory of general relativity – the expressions of all of the laws of nature in all possible continuously connected frames of reference, from the view of any one of them, must be in *one-to-one correspondence*. The different reference frames, in turn, relate to each other in terms of continuous, differentiable transformations, that we call "motion". [The differentiability of these transformations to all orders, requiring that they be analytic, is dictated by the requirement of the inclusion of the laws of conservation of energy, momentum and angular momentum in the special relativity limit of the theory; this is in accordance with Noether's theorem[7]] When the relative motion is inertial, characterized by 3 constant parameters of relative velocity, 3 (Eulerian) angles of rotation and 4 translations, the underlying set of transformations forms the 10-parameter Lie group of special relativity. It is the Poincare group **P**.[2] This a special limit of the coordinate-dependent (non-inertial) transformations group of general relativity - the Einstein group **E**. It is a 16-parameter Lie group whose representations are a global extension of the representations of the Lie group **P**.[21]

It is important to recognize that, in physics, **P** is an *asymptotic limit* of **E**, but **P** is *not* a subgroup of **E**. This is for the physical reason that **P** is only exact in the case of a vacuum – wherein the entire universe would be empty! For in the field theory of general relativity, if there should be matter, *anywhere in the universe,* the continuous, analytic fields associated with this matter must be nonzero *everywhere.* In this case, the parameters that relate a reference frame to any other must be spacetime-dependent. Thus, special relativity can only be viewed as an asymptotic limit of general relativity. Its representations may be approached asymptotically from those of general relativity, as closely as we please, *but not reached in an exact sense!* The Einstein Group **E** is a form of a topological group **T**.

A Mathematical diversion on the Nature of **E** – *Pontjagin's Theorem*

Firstly, because **E** prescribes the invariance associated with continuous changes from any point of the function space of field solutions, to any other that is arbitrarily close, **E** must be *locally compact.*

Secondly, because of the rejection of the discontinuous reflections in the spacetime, the topological space of this group must be *connected* – i.e. it cannot be decomposed into two or more disjoint sets.

Thirdly, since the elements of this topological space are a countable number of fields $\{\psi^{(1)}(x), \ldots \psi^{(n)}(x)\}$ – corresponding to the countable modes of the closed system – and since the continuous changes of these field variables in their own neighborhoods $\{\delta\psi^{(1)}, \ldots \delta\psi^{(n)}\}$ are induced by the transformations of the group, it follows that the complete set of neighborhoods of the topological space is *countable.* The topological group **T** is then said to satisfy the *second axiom of countability.*

Pontrjagin's theorem is the following:[22] Let **T** be a locally compact, connected topological field satisfying the second axiom of countability. Then **T** is isomorphic with one of the three topological fields: a) the field of real numbers, b) the field of complex numbers, c) the field of quaternions.

Since the Einstein group **E** corresponds with the topological group **T**, the most general mathematical system to express the laws of physics in general relativity is then the set of quaternions.[23] Reducing the quaternion-valued field from four dimensions to two leads to the complex number-valued field and further reduction to one dimension leads to the real number-valued field. [The word "field" here refers to an "algebraic field"[22]]. The latter two sets may be seen as subsets of the first, reductions where one loses not only dimensionality but also the important feature of noncommutability of the quaternion number system. The quaternion field then expresses the laws of nature to be compatible with the covariance requirement of the group of general relativity **E**.

The Electromagnetic Field Equations in General Relativity

The vector representation (2.5) of Maxwell's equations extends to general relativity by globalizing the ordinary derivatives to covariant derivatives that entail the affine connection $\Gamma^{\lambda}{}_{\mu\nu}$ of the curved spacetime.[24] Thus, (2.5) takes the following form in the curved spacetime:

$$F_{\mu\nu}{}^{;\nu} = 4\pi j_{\mu} \qquad (4.1a)$$

$$F^{[\mu\nu}{}_{;\lambda]} = 0 \qquad (4.1b)$$

Where the square brackets denote the cyclic sum and

$$F^{[\alpha\beta}{}_{;\gamma]} \equiv \partial_{\gamma}F^{\alpha\beta} + \Gamma^{\alpha}{}_{\rho\gamma}F^{\rho\beta} + \Gamma^{\beta}{}_{\rho\gamma}F^{\alpha\rho} \qquad (4.2)$$

The affine connection coefficients in terms of the metric tensor are:[24]

$$\Gamma^{\rho}{}_{\mu\alpha} = (1/2)g^{\rho\lambda}(\partial_{\mu}\, g_{\lambda\alpha} + \partial_{\alpha}\, g_{\mu\lambda} - \partial_{\lambda}\, g_{\alpha\mu})$$

The two-component spinor form of electromagnetic field theory (3.7) is generalized in the curved spacetime by 1) globally extending the Pauli matrices to the quaternion elements, $\sigma^{\mu} \rightarrow q^{\mu}(x)$, and b) generalizing the ordinary derivatives to covariant derivatives of the spinor variables. This entails the 'spin-affine connection fields' Ω_{μ} as follows:

$$q^{\mu}(x)\phi_{\alpha;\mu} \equiv q^{\mu}(x)(\partial_{\mu} + \Omega_{\mu})\phi_{\alpha} = \Upsilon_{\alpha} \qquad (4.3)$$

where

$$\Omega_{\mu} = (1/4)(\partial_{\mu}q^{\rho} + \Gamma^{\rho}{}_{\tau\mu}q^{\tau})q_{\rho}{}^{*} \qquad (4.4)$$

and $q_{\rho}{}^{*}$ is the quaternion conjugate to q_{ρ}, corresponding to time-reversal (or space reflection) of q_{ρ}. [Henceforth, the asterisk over the quaternion variables denotes the 'quaternion conjugate', not the 'complex conjugate'. The former is obtained by reversing the sign of the time component of the quaternion variable].

The Global Spinor Lagrangian for Electromagnetism

The Lagrangian density that gives, upon variation, the topologically covariant field equations (4.3) is an explicit function of the spinor variables, ϕ_1, $\phi_1{}^{+}$, ϕ_2, $\phi_2{}^{+}$ and their respective covariant derivatives. [The superscript 'dagger' denotes the Hermitian conjugate of the function]. It has the form:

$$L_M = \{ig_M \Sigma_\alpha (-1)^\alpha [\phi_\alpha^+ (q^\mu \phi_{\alpha;\mu} - 2\Upsilon_\alpha) + \text{h.c.}\} (-g)^{1/2} \quad (4.5)$$

where 'h.c.' stands for the Hermitian conjugate of the term preceding it and

$$(-g)^{1/2} = i\varepsilon_{\mu\nu\lambda\rho} q^\mu q^{\nu*} q^\lambda q^{\rho*}$$

is the 'metric density'. The multiplicative constant g_M in (4.5) has the dimension of a *length* – it is the one extra fundamental constant in this theory. Its appearance results from the generalization that is effected when the Lagrangian is expressed in terms of the spinor variables rather than the usual vector variables. Since the spinor variables ϕ_α have the dimension of an electric field intensity, the terms summed have a dimension of energy density per length; thus g_M has the dimension of a length so that the Lagrangian has the proper dimension of energy density.

The magnitude of g_M was determined from the prediction that this spinor formalism yields the Lamb shift in hydrogenic atoms. It follows from the new terms in the spinor formulation of electromagnetism that appear in the Dirac Hamiltonian, that are not present in the standard Dirac equation for hydrogen.[16] These extra terms then predict a lifting of the accidental degeneracy in the states of hydrogen, thus the Lamb shift. The new fundamental constant g_M was found to have a magnitude that is the order of 2×10^{-14} cm. This gives results that are in agreement with the experimental facts on the Lamb shifts $nS_{1/2} - nP_{1/2}$ for the principal quantum numbers for hydrogen, n = 2, 3 and 4.

Derivation of the Maxwell Field Formalism from General Relativity

The covariance groups underlying the tensor forms of the respective Einstein and the Maxwell field equations are reducible. This is because they entail reflection symmetry, not required by relativity theory, as well as the required continuous symmetry of the Einstein group **E.** When the Einstein field equations are factorized, they yield the irreducible form, that are then in terms of the quaternion and spinor variables, rather than the tensor variables. Such a generalization must then extend the physical predictions of the usual tensor forms of general relativity of gravitation and the standard vector representation of the Maxwell theory (both in terms of second-rank tensor fields, one symmetric and the other antisymmetric) because the new factorized variables have more degrees of freedom than the earlier version.

The starting point then to achieve the factorization of the Einstein equations is the factorized differential line element in the quaternion form, $ds = q^\mu(x)dx_\mu$, where q^μ are a set of four quaternion-valued components of a four vector. Thus ds is, geometrically, a scalar invariant, but it is algebraically a quaternion. As such, it behaves like a second rank spinor of the type $\psi \otimes \psi^*$, where ψ is a two-component spinor variable.[17]

We see then that the basic variable that represents the generalized spacetime that is appropriate to general relativity is a 16-component variable. One may then speculate at the outset that these 16 independent components of the metrical field relate to the 10 components of the gravitational field plus 6 components of the Maxwell field, in terms of a single unified field that incorporates both gravitation and electromagnetism. We will now see that this is indeed the case.

Since there is no reflection symmetry in the quaternion formulation, the 'reflected' quaternion $q^{\mu *}$ must be distinguishable from q^{μ}. The conjugate differential line element to ds is $ds^* = q^{\mu *}dx_{\mu}$. The product of the quaternion and conjugate quaternion line elements is then the real number-valued element that corresponds to the squared differential element of the Riemannian geometry:

$$ds\, ds^* = -(1/2)(q^{\mu} q^{\nu *} + q^{\nu}q^{\mu *})dx_{\mu}\, dx_{\nu} \Leftrightarrow \sigma_0 g^{\mu\nu}dx_{\mu}\, dx_{\nu}$$

Thus the symmetric second rank metric tensor $g^{\mu\nu}$ of Einstein's formulation of general relativity corresponds to the symmetric sum from the quaternion theory, $(-1/2)(q^{\mu}q^{\nu *} + q^{\nu}q^{\mu *})$. [The factor (-1/2) is chosen in anticipation of the normalization of the quaternion variables.] Thus we see that ds is a factorization of the standard Riemannian squared differential metric $ds^2 = g^{\mu\nu}dx_{\mu}dx_{\nu}$.

The following is an outline that leads to a derivation of the factorization of the Einstein formalism that gives back to the gravitational and the electromagnetic equations from a unified quaternion-spinor formalism.

The Variables of a Riemannian Spacetime in Quaternion Form[25]

Let us now exploit the feature of the quaternion metrical field that it has configuration degrees of freedom, as a four-vector, as well as spinor degrees of freedom, as a second rank spinor of the type: $\psi\otimes\psi^*$.

Since the quaternion q^{μ} is a four-vector, the product $q^{\mu}q_{\mu}^*$ must be invariant under the continuous spacetime transformations *and* the reflections in spacetime. It then follows that the covariant derivatives and the second covariant derivatives of the quaternion fields must vanish. Since $q^{\mu} \sim (\psi\otimes\psi^*)^{\mu}$, it follows that (with $\alpha,\beta = 1,2$)

$$0 = (q_{\mu;\rho;\lambda} - q_{\mu;\lambda;\rho})_{\alpha\beta} = [(\psi_{\alpha;\rho;\lambda} - \psi_{\alpha;\lambda;\rho})\psi_{\beta}^* + \psi_{\alpha}(\psi_{\beta;\rho;\lambda} - \psi_{\beta;\lambda;\rho})] + ([q_{\mu;\rho;\lambda}] - [q_{\mu;\lambda;\rho}]) \qquad (4.6)$$

The squared bracket above denotes the behavior of the quaternion field with respect to its vector degrees of freedom alone. The covariant derivatives of the two-component spinor variables are as follows: $\psi_{;\rho} = (\partial_{\rho}$

$+ \ \Omega_\rho)\psi$ and the "spin affine connection" has two alternative (equivalent) forms:[17]

$$\Omega_\rho = (1/4)(\partial_\rho q^\mu + \Gamma^\mu{}_{\tau\mu} q^\tau) q_\mu{}^* = -(1/4) q_\mu (\partial_\rho q^{\mu *} + \Gamma^\mu{}_{\tau\rho} q^{\tau *}) \quad (4.7)$$

Ω_ρ is the term that must be added to the ordinary derivative of a spinor field in a curved spacetime in order to define its derivative covariantly; that is to say, in order that the spinor variable is integrable in the curved spacetime.

The first two terms on the right hand side of eq. (4.6) denote the changes with respect to the spinor indices, the third term denotes the changes in configuration space. Their explicit forms are:

$$\psi_{;\rho;\lambda} - \psi_{;\lambda;\rho} = (\partial_\lambda \Omega_\rho + \Omega_\lambda \Omega_\rho - \partial_\rho \Omega_\lambda - \Omega_\rho \Omega_\lambda)\psi \equiv K_{\lambda\rho}\psi \quad (4.8)$$

where $K_{\rho\lambda} = -K_{\lambda\rho}$ is the "spin curvature tensor". It is clearly a second-rank, antisymmetric tensor in configuration space. Since the left-hand side of this equation is a first-rank spinor in spinor space, the spin curvature tensor on the right, $K_{\rho\lambda,}$ must be a second rank spinor that contracts with the first rank spinor ψ on the right to yield a first-rank spinor function. The spin affine connection field Ω_ρ, on the other hand, is a four-vector in configuration space, but it is not covariant in spinor space. This is clear since it is the term that must be added to the ordinary (non-covariant) derivative of the spinor variable in order to make its derivative in a curved spacetime covariant.

In the third term in eq. (4.6), where the square bracket stands for the changes in q_μ as a four vector, we have:

$$[q_{\mu;\rho;\lambda}] - [q_{\mu;\lambda;\rho}] = R_{\kappa\mu\rho\lambda} q^\kappa \quad (4.9)$$

This defines the "Riemann curvature tensor", $R_{\kappa\mu\rho\lambda}$, wherein q_μ could be any covariant four-vector.

Substituting (4.8) and (4.9) into (4.6), the relation between the spin curvature tensor and the Riemann curvature tensor follows:

$$K_{\rho\lambda} q_\mu + q_\mu K_{\rho\lambda}{}^+ = -R_{\kappa\mu\rho\lambda} q^\kappa \quad (4.10)$$

where the 'dagger' denotes the Hermitian adjoint of the function.

In a similar fashion, application of the preceding analysis to the conjugated quaternion fields yields the accompanying equation to (4.10):

$$K_{\rho\lambda}{}^+ q_\mu{}^* + q_\mu{}^* K_{\rho\lambda} = R_{\kappa\mu\rho\lambda} q^{\kappa *} \quad (4.11)$$

Multiplying (4.10) on the right with the conjugated quaternion $q_\gamma{}^*$ and (4.11) on the left with the quaternion q_γ, then adding the resulting equations and using the identity:

$$q_\gamma q^{\kappa *} + q^\kappa q_\gamma{}^* = 2\sigma_0 \delta_\gamma{}^\kappa$$

where σ_0 is the unit two-dimensional matrix, the following correspondence is derived between the Riemann curvature tensor and the spin curvature tensor:

$$\sigma_0 R_{\kappa\mu\rho\lambda} = (1/2)(K_{\rho\lambda}q_\mu q_\kappa^* - q_\kappa q_\mu^* K_{\rho\lambda} + q_\mu K_{\rho\lambda}^+ q_\kappa^* - q_\kappa K_{\rho\lambda}^+ q_\mu^*) \quad (4.12)$$

Next, contracting $R_{\kappa\mu\rho\lambda}$ with the contravariant metric tensor $g^{\mu\lambda}$ yields the correspondence with the Ricci tensor $R_{\kappa\rho}$:

$$\sigma_0 g^{\mu\lambda} R_{\mu\kappa\rho\lambda} \equiv \sigma_0 R_{\kappa\rho} = (1/2)(K_{\rho\lambda}q^\lambda q_\kappa^* - q_\kappa q^{\lambda *} K_{\rho\lambda} + q^\lambda K_{\rho\lambda}^+ q_\kappa^* - q_\kappa K_{\rho\lambda}^+ q^{\lambda *}) \quad (4.13)$$

Finally, the scalar curvature field R follows from the further contraction of the Ricci tensor (4.13) with the metric tensor, giving:

$$\sigma_0 R = (1/2)(K_{\rho\lambda}q^\lambda q^{\rho *} - q^\rho q^{\lambda *} K_{\rho\lambda} + q^\lambda K_{\rho\lambda}^+ q^{\rho *} - q^\rho K_{\rho\lambda}^+ q^{\lambda *}) \quad (4.14)$$

The Quaternion Field Equations

The Lagrangian density whose vanishing variation leads to the field equations in q^λ is chosen to be the trace of the scalar curvature:

$$L_E = (\mathrm{Tr}R)(-g)^{1/2} = (1/2)\mathrm{Tr}(q^{\rho *} K_{\rho\lambda} q^\lambda + \text{h.c.})(-g)^{1/2} \quad (4.15)$$

If we signify by L_M the part of the Lagrangian density that yields the matter variables upon variation with respect to the quaternion variables, then the total Lagrangian density is $L = L_E + L_M$. Its variation with respect to the conjugated quaternion variables then yields the field equations:[25]

$$(1/4)(K_{\rho\lambda}q^\lambda + q^\lambda K_{\rho\lambda}^+) + (1/8)Rq_\rho = kT_\rho \quad (4.16a)$$

Variation with respect to the quaternion variables yields the conjugated quaternion field equation:

$$-(1/4)(K_{\rho\lambda}^+ q^{\lambda *} + q^{\lambda *} K_{\rho\lambda}) + (1/8)Rq_\rho^* = kT_\rho^* \quad (4.16b)$$

[Note that the source term T_ρ is a quaternion and T_ρ^* is a conjugated quaternion].

The quaternion field equations (4.16ab) are then the factorization of Einstein's tensor field equations:

$$R_{\mu\nu} - (1/2)g_{\mu\nu}R = kT_{\mu\nu} \quad (4.17)$$

The solutions of the latter equations are the ten components of the symmetric second-rank metric tensor $g_{\mu\nu}$. The solutions of the factorized equations 4.16a (or 4.16b) are the sixteen components of the quaternion metrical field q_ρ (or q_ρ^*). We will now see that this sixteen-component metrical quaternion field indeed incorporates the gravitational and the electromagnetic fields in terms of their earlier tensor representations. Gravitation entails ten of the components in the symmetric second-rank tensor $g_{\mu\nu}$. Electromagnetism entails six of the components (the three components of the electric field and the three components of the magnetic field), as incorporated in the second-rank antisymmetric tensor $F_{\mu\nu}$.

To demonstrate the natural unification of the gravitational and electromagnetic aspects of the quaternion field equation (4.16a) and its conjugate equation (4.16b), we follow this procedure: Multiply (4.16a) on the right with the conjugated quaternion solution q_γ^*, and the conjugated equation (4.16b) on the left with q_γ. Then adding (with the constant k on the right) and subtracting (with the constant k' on the right) we obtain the following pair of equations:

$$(1/2)(K_{\rho\lambda}q^\lambda q_\gamma^* -(\pm)\ q_\gamma q^{\lambda *}K_{\rho\lambda} + q^\lambda K_{\rho\lambda}{}^+ q_\gamma^* -(\pm)\ q_\gamma K_{\rho\lambda}{}^+ q^{\lambda *}) + (1/4)(q_\rho q_\gamma^* \pm q_\gamma q_\rho^*)R = 2\ (T_\rho q_\gamma^* \pm q_\gamma T_\rho^*) \qquad (4.18\pm)$$

Examination of eqs. (4.12), (4.13) and (4.14) shows that eq. (4.18+) is in one-to-one correspondence with Einstein's second-rank symmetric tensor equation (4.17).

The Electromagnetic Field Equations

The antisymmetric second-rank tensor equations (4.18-), corresponding with six independent relations, may be re-expressed in terms of the Maxwell field theory (2.5) by taking the covariant divergence of (4.18-), with

$$F_{\rho\gamma} = Q[(1/4)(K_{\rho\lambda}q^\lambda q_\gamma^* + q_\gamma q^{\lambda *}K_{\rho\lambda} + q^\lambda K_{\rho\lambda}{}^+ q_\gamma^* + q_\gamma K_{\rho\lambda}{}^+ q^{\lambda *}) \quad (1/8)(q_\rho q_\gamma^* - q_\gamma q_\rho^*)R] \qquad (4.19)$$

In this expression for the electromagnetic field intensity tensor, Q is a constant of proportionality with the dimension of charge, inserted on both sides of eq. (4.18-). The four-current density is:

$$j_\gamma = (Qk'/4\pi)(T_\rho{}^{;\rho}q_\gamma^* - q_\gamma T_\rho{}^{;\rho *}) \qquad (4.20)$$

The role of the *Mach principle* is revealed at this stage of the analysis. Since $F_{\rho\lambda}$ depends on the spin curvature tensor $K_{\rho\lambda}$, which automatically vanishes in a vacuum (i.e. a flat spacetime), the electromagnetic field, and therefore the previously considered electric charge of any quantity of matter in a vacuum must vanish. Thus, not only the inertial mass but also the electric charge of a 'particle' of matter does not exist when there is no coupling to other matter. I have generalized this idea in the field theory based on General Relativity, to the case where *all* previously considered *intrinsic properties* of discrete matter, in addition to inertial mass and electric charge, vanish identically in a vacuum. This view exorcises all of the remaining features of the discrete, separable 'elementary particle' of matter. It is replaced with a view of matter in terms of a closed, continuous field theory, according to the theory of general relativity. I have called this view of matter, whereby all of its previously considered intrinsic properties are explained in terms of coupling within the closed system, '*the generalized Mach principle*'.[10]

The Conservation of Charge

Since $F_{\rho\gamma}$ is an antisymmetric tensor in spacetime and since the components of the ordinary affine connection $\Gamma_{\alpha\beta}{}^{\gamma}$ are symmetric in the indices $(\alpha\beta)$, it follows that the four-divergence of the current density j_γ automatically vanishes. That is, as in the standard formulation, the equation of continuity follows from taking the covariant divergence of Maxwell's equation (4.1a):

$$j_\gamma{}^{;\gamma} = (1/4\pi)F_{\rho\gamma}{}^{;\rho;\gamma} = 0 \qquad (4.21)$$

It then follows from the integral form of (4.21) (in the local domain) that the integral of the time component j_0 over all of 3-space is time-conserved. This assumes that there is no current flow in or out of the surface containing the charge $Q = \int j_0 d^3x$, that gives rise to the electromagnetic field of force $F_{\rho\gamma}$.

The Absence of Magnetic Monopoles[26]

The form (4.18-) for the electromagnetic field intensity was seen to yield four out of the eight of Maxwell's equations associated with the current source, as shown in eq (4.1a). It also follows that the four of Maxwell's equations without source terms,

$$F^{[\mu\nu}{}_{;\lambda]} \equiv F^{\mu\nu}{}_{;\lambda} + F^{\lambda\mu}{}_{;\nu} + F^{\nu\lambda}{}_{;\mu} = 0 \qquad (4.22)$$

are predicted by the quaternion structure of $F_{\rho\gamma}$ as given in (4.19). This implies the absence of magnetic monopoles since, if they did exist, the right-hand side of (4.22) would be non-zero.

This result is a consequence of the dependence of the definition of $F_{\rho\gamma}$ of the spin curvature tensor, $K_{\rho\lambda}$, according to eq. (4.19). It is because the spin curvature tensor $K_{\rho\lambda}$ is the four-dimensional curl of a four-vector in configuration space:

$$K_{\rho\lambda} = \partial_\lambda\Omega_\rho - \partial_\rho\Omega_\lambda + \Omega_\lambda\Omega_\rho - \Omega_\rho\Omega_\lambda = \Omega_{\rho;\lambda} - \Omega_{\lambda;\rho} = -K_{\lambda\rho} \quad (4.23)$$

which, in turn, follows from the transformation of the spin affine connection Ω_ρ *in configuration space* as a four-vector. It then follows that the cyclic sum $K_{[\rho\lambda;\gamma]} = 0$.

This result, according to eq. (4.19), in turn, implies that eq. (4.22), $F_{[\rho\lambda;\gamma]} = 0$, must be true, indicating that there are no magnetic monopoles in this formulation of the electromagnetic field equations – for if there were, there would be a non-zero source term in eq. (4.22).

Thus we have seen that the factorized quaternion field equations (4.16a) (or their conjugated equations 4.16b) –the irreducible form of electromagnetism according to the underlying group of general relativity – indicates a lack of magnetic monopoles, in agreement with the standard formulation of the Maxwell field theory. The factorized field formulation of general relativity, in terms of the 16-component quaternion metrical field, q^{μ}, then automatically fuses the laws according to the Einstein's formulation, and the laws of electromagnetism, according to the Maxwell formulation, in a *unified field theory* of the gravitational and electromagnetic manifestations of matter.

> *It is important to recognize at this stage of the analysis that the unified field equations (4.16) entail more physical predictions than do the respective earlier versions of gravitation – Einstein's field equations– and electromagnetism – Maxwell's field equations. We have already seen extra predictions from the spinor form of the electromagnetic field equations. Though this theoretical analysis does not focus on the expression of electromagnetism in terms of potential fields, extra predictions may indeed also follow in the quaternion-spinor formulation from the structuring of the electromagnetic $B^{(3)}$ potential field, as derived in the theory of M. Evans and his collaborators.*[27]

5.Electromagnetism and Wave Mechanics[28]

Derivation of Born's Probability Calculus

The conventional conceptual content of Quantum Mechanics was initiated by the Copenhagen School when it was recognized that one could express the linear Schrodinger wave mechanics in terms of a probability calculus, whose solutions are represented with a Hilbert function space. Max Born then interpreted the wave nature of matter in terms of a spatially distributed *probability amplitude* - a wave represented by a complex function – to accompany the material particle as it moves from one place to another. The Copenhagen view was then to *define* the basic nature of matter in terms of the measurement process, with an underlying probability calculus - wherein the probability densities (for locating the particles of matter/volume) are the real number-valued modulii of the matter wave amplitudes.

But this was not Schrodinger's intention in his formulation of wave mechanics![29] Rather, it was to complete the Maxwell field formulation of electromagnetic theory by incorporating the empirically verified wave nature of matter in the source terms on the right hand sides of Maxwell's equations.

The 'matter field' was originally postulated by Louis de Broglie, and discovered in the electron diffraction studies of Davisson and Germer[30] and of G.P. Thomson[31]. From Schrodinger's understanding of the matter field of, say, an electron, it must be represented in the source terms (charge and current density) of Maxwell's equations, as the modulii of these waves.

Integration of the *local limit* of eq. (4.20) for the four-current density source of Maxwell's equations, together with the boundary condition that the 3-current part of this density, j_k, vanishes on the bounding surface of a volume that contains the total charge Q, then gives the law of conservation of electric charge:

$$\sigma^0 Q = \int j_0 d^3x = \textit{constant in time} \qquad (5.1)$$

[The insertion of the unit matrix σ^0 on the left takes account of the matrix structure (3.15) of the current density in the quaternion formulation.]

The local limit of j_0 in eq. (4.20) is:

$$j_0 = -(Qk'/4\pi)\partial^\rho(T_\rho^{(1)} + T_\rho^{(1)*}) \qquad (5.2)$$

where $T_\rho^{(1)}$ is the local limit of the matter source of the quaternion metric field equation. In terms of the Lagrangian density for the matter variables, its form is $-\delta L_M/\delta q^{\rho*}$.

Taking determinants of both sides of eq. (5.2) then yields the value of the constant k' as follows:

$$k' = -(4\pi)/\left|\int\partial^\rho(T_\rho^{(1)} + T_\rho^{(1)*})d^3x\right|$$

where the 'vertical bars' denote the determinant.

Thus, the four-current density j_γ of this expression of the Maxwell theory has the following general form in a curved spacetime:

$$j_\gamma = Q(T_\rho{}^{;\rho}q_\gamma - q_\gamma T_\rho{}^{;\rho*})/\left|\int\partial^\rho(T_\rho^{(1)} + T_\rho^{(1)*})d^3x\right| \qquad (5.3)$$

The 'matter density', interpreted in conventional Quantum Mechanics as a 'probability density', is then:

$$\sigma^0\rho = j_0/Q = (q_0 T_\rho{}^{;\rho*} - T_\rho{}^{;\rho}q_0^*)/\left|\int\partial^\rho(T_\rho^{(1)} + T_\rho^{(1)*})d^3x\right| \qquad (5.4)$$

In the local limit, $q_0 \rightarrow \sigma_0$, $q_0^* \rightarrow \sigma_0^* = -\sigma_0$, $T_\rho{}^{;\rho} \rightarrow \partial^\rho T_\rho^{(1)}$

Thus, we have:

$$\text{loc lim}\left|\int\rho d^3x\right| = 1 \qquad (5.5)$$

Equation 5.5 is the normalization condition that was *postulated* by Max Born, in his interpretation of Schrodinger's non-relativistic wave mechanics as a probability calculus. As we see here, the *derived* normalization is not a general relation in the full, generally covariant expression of the field theory.

We also see from the general form (5.3) that the three-current density part of j_γ is:

$$j_k/Q = (q_k T_\rho{}^{;\rho *} - T_\rho{}^{;\rho} q_k^*) / \left| \int \partial^\rho (T_\rho{}^{(1)} + T_\rho{}^{(1)*} d^3x \right|$$

This expression predicts a coupling of the 'gravitational field' (in terms of q_k) with the matter field components T_ρ to define a gravitational current contribution. The latter is not foreseen in the conventional theories that neglect the gravitational coupling to matter fields.

Summary

We have seen in this section that the factorization of Einstein's symmetric, second-rank tensor field equations (10 relations) to a quaternion form (16 relations) yields not only the gravitational and electromagnetic manifestations of matter in a unified field theory, but that they also reveal a feature of Quantum Mechanics. In particular, it was found that in the flat space approximation to the curved space representation in general relativity, the time component of the electromagnetic four-current density corresponds in a one-to-one way with the probability density of Quantum Mechanics. Its integration over all of space in this limit is found to be unity.

This is a result that was postulated (not derived from first principles) when Born attempted to identify Quantum Mechanics with a probability calculus. The result of this analysis, in which the normalization follows as a derivation from General Relativity, together with a rigorous derivation of the quantum mechanical equations from general relativity[32] then enforces the view of a paradigm change in physics. It is from that of Quantum Theory, which has dominated the last two thirds of the twentieth century, to that of General Relativity, as a theory of electromagnetism, gravity and matter, in all domains. This is a shift to a paradigm for the laws of matter based fully on the views of continuity, determinism and holism, in terms of the nonsingular field concept.

BIBLIOGRAPHY

1. A. Einstein, *The Meaning of Relativity* (Princeton, 1955), Fifth Edition. English translations of some of the initial papers on relativity theory are in: A. Einstein, H.A. Lorentz, H. Weyl and H. Minkowski, *The Principle of Relativity* (Dover, 1923).

2. Y.S. Kim and M.E. Noz, *Theory and Applications of the Poincare Group* (Reidel, 1986). One of the pioneering books on group theory that clearly explicates the spinor role in relativistic covariance is: E.P. Wigner, *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra* (Academic, 1959).

3. M. Sachs, *Relativity in Our Time* (Taylor and Francis, 1993).

4. The mathematical details of the approach of Einstein's general relativity in approaching Newton's theory of universal gravitation are outlined in: R.J. Adler, M.J. Bazin and M. Schiffer, *Introduction to General Relativity* (McGraw-Hill, 1975) Second Edition, Chapter 10. Further discussion is in ref. 3, Chapter 19.

5. A. Einstein, *Annals of Mathematics* **46,** 578 (1945).

6. L.P. Eisenhart, *Continuous Groups of Transformations* (Dover, 1961).

7. Noether's theorem is derived in: E. Noether, "Invariante Variationsprobleme", *Goett. Nachr* (1918), 235. It is explicated further in: C. Lanczos, *The Variational Principles of Mechanics* (Toronto, 1966) Third Edition, Appendix II and in: R. Courant and D. Hilbert, *Mathematical Methods of Physics* (Interscience 1953) Volume I, p. 262. Application to fields is clearly demonstrated in N.N. Bogoliubov and D.V. Shirkov, *Introduction to the Theory of Quantized Fields* (Interscience, 1980) Third Edition, Sec. 1.2.

8. In *James Clerk Maxwell: A Commemorative Volume 1831 – 1931*, ed. J. J. Thomson, (Cambridge, 1931), Einstein said: "I incline to the belief that the physicists will … be brought back to the programme which may suitably be called Maxwell's: the description of Physical Reality by fields which satisfy without singularity a set of Partial differential equations". In a letter that Einstein wrote to David Bohm in 1953, he said: "If it is not correct that reality is described as a continuous field, then all my efforts are futile", *Einstein Archives*, Jewish National and University Library, The Hebrew University of Jerusalem, Call No. 1576:8-053.

9. A complete discussion of Lie groups is in: H. Weyl, *The Theory of Groups in Quantum Mechanics* (Dover, 1931) transl. H. Robertson, Section 15.

10. I have discussed the 'Generalized Mach Principle' in: M. Sachs, *General Relativity and Matter* (Reidel, 1982), Sec. 1.7.

11. C.S. Wu, E. Ambler, R.W. Hayward, D.D. Hoppes and R.P. Hudson, *Phys. Rev.* **105,** 1413 (1957).

12. M. Sachs, *Annals of Physics* **6,** 244 (1959).

13. M. Sachs, *General Relativity and Matter* (Reidel, 1982) Sec. 4.7.

14. M. Sachs and S.L. Schwebel, *Annals of Physics* **8**, 475 (1959).

15. N.F. Ramsey, *Physics Reports* **43**, 409 (1978).

16. M. Sachs, *Quantum Mechanics from General Relativity* (Reidel, 1986) Sec. 8.2.

17. Ref. 13, chapter 3.

18. *Ibid.* chapter 5.

19. *Ibid.*, equation (3.31).

20. *M.* Sachs, "Relativistic Implications of Electromagnetic Field Theory", in T.W. Barrett and D.M. Grimes, eds. *Advanced Electromagnetism*(World Scientific, 1995), p. 551.

21. *M.* Sachs, *Lett. Nuovo Cimento* **21,** 123 (1978).

22. L. Pontrjagin, *Topological Groups* (Princeton, 1939), p. 173.

23. In a letter to P.G. Tait in 1871, J.C. Maxwell said the following about the use of quaternions in the laws of physics: "..the virtue of the 4nions lies not so much as yet in solving hard questions as in enabling us to see the meaning of the question and its solutions.." *Archives,* Cavendish Laboratory, Cambridge.

24. For a derivation of the affine connection in the tensor formalism, see ref. 4.

25. Ref. 13, chapter 6.

26. M. Sachs, *Nuovo Cimento* **114B**, 123 (1999).

27. M.W. Evans and J.-P. Vigier, *The Enigmatic Photon, Volume1 : The Field $B^{(3)}$* (Kluwer, 1994); M.W. Evans, J.-P. Vigier, S. Roy and S. Jeffers, *The Enigmatic Photon, Volume3: Theory and Practice of the $B^{(3)}$ Field* (Kluwer, 1996).

28. M. Sachs, *Annales de la Fondation Louis de Broglie* **17**, 163 (1992).

29. M. Sachs, *Einstein Versus Bohr* (Open Court, 1988), Chapter 5.

30. C.J. Davisson and L.H. Germer, *Phys. Rev.* **30,** 705 (1927).

31. G.P. Thomson, *Proc. Roy. Soc.* (London) **A117**, 600 (1928).

32. M. Sachs, *General Relativity and Matter* (Reidel, 1982), Chapter 4.